The Dark Energy Survey as a Large Data Set

Josh Frieman
Fermilab and University of Chicago

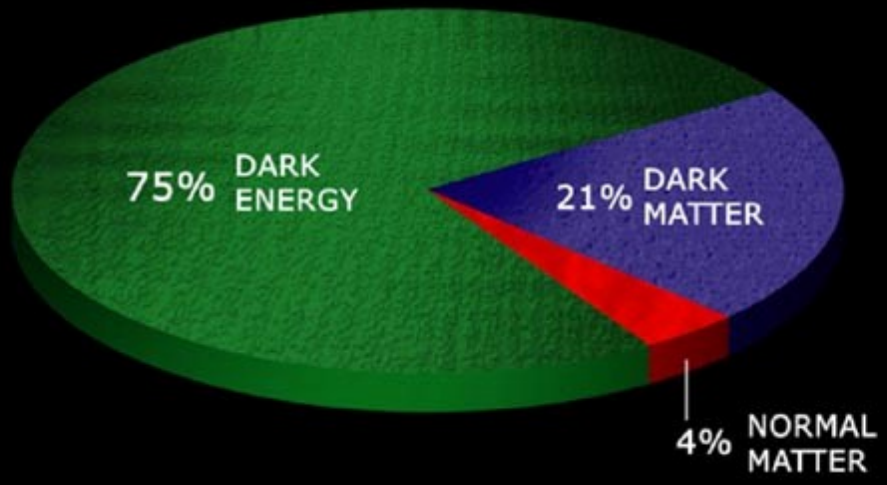
ICiS Workshop on
Large Data Sets in Astrophysics and Cosmology
August 2011

www.darkenergysurvey.org

DARK ENERGY SURVEY

Goal: Probe Dark Energy

- What is the physical cause of cosmic acceleration?
  - Dark Energy or modification of General Relativity?
    - If Dark Energy, is it Λ (the vacuum) or something else?
      - What is the DE equation of state parameter *w*?



The Dark Energy Survey

- Survey project using 4 complementary techniques:
  - I. Cluster Counts
  - II. Weak Lensing
  - III. Large-scale Structure
  - IV. Supernovae
- Two multiband surveys:
  - 5000 deg^2 *grizY* to 24th mag
  - 30 deg^2 repeat (supernovae)
- Build new 3 deg^2 FOV camera and Data management system
  - Survey 2012-2017 (525 nights)
  - Facility instrument for Blanco

Blanco 4-meter at CTIO

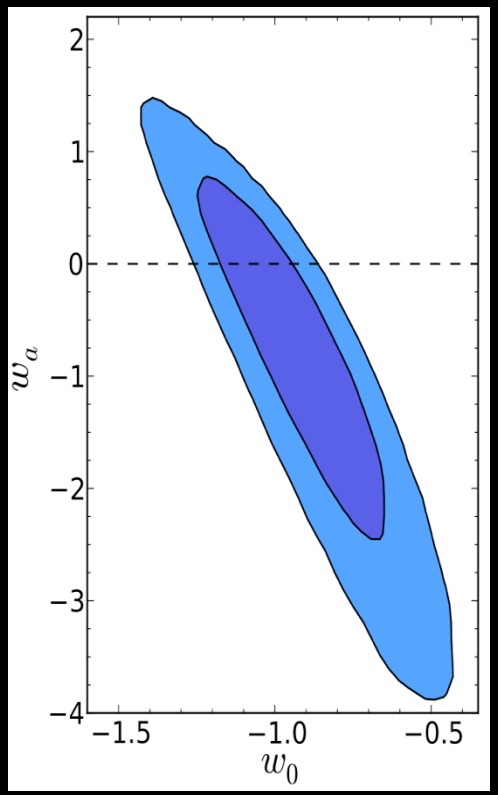
DES Science Summary

DARK ENERGY SURVEY

Four Probes of Dark Energy

- **Galaxy Clusters**
  - ~100,000 clusters to z>1
  - Synergy with SPT, VHS
  - Sensitive to growth of structure and geometry
- **Weak Lensing**
  - Shape measurements of 300 million galaxies
  - Sensitive to growth of structure and geometry
- **Baryon Acoustic Oscillations**
  - 300 million galaxies to z = 1 and beyond
  - Sensitive to geometry
- **Supernovae**
  - 30 sq deg time-domain survey
  - ~4000 well-sampled SNe Ia to z ~1
  - Sensitive to geometry

Current Constraints on DE Equation of State



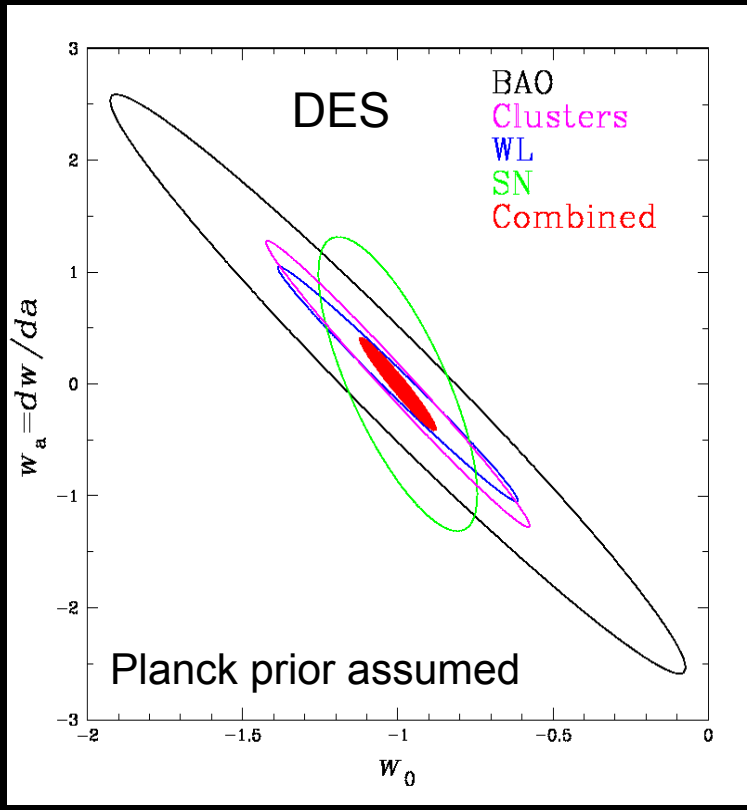
DES Science Summary

DARK ENERGY SURVEY

Four Probes of Dark Energy

- **Galaxy Clusters**
  - ~100,000 clusters to z>1
  - Synergy with SPT, VHS
  - Sensitive to growth of structure and geometry
- **Weak Lensing**
  - Shape measurements of 300 million galaxies
  - Sensitive to growth of structure and geometry
- **Baryon Acoustic Oscillations**
  - 300 million galaxies to z = 1 and beyond
  - Sensitive to geometry
- **Supernovae**
  - 30 sq deg time-domain survey
  - ~4000 well-sampled SNe Ia to z ~1
  - Sensitive to geometry

Forecast Constraints on DE Equation of State



Factor 3-5 improvement over Stage II DETF Figure of Merit

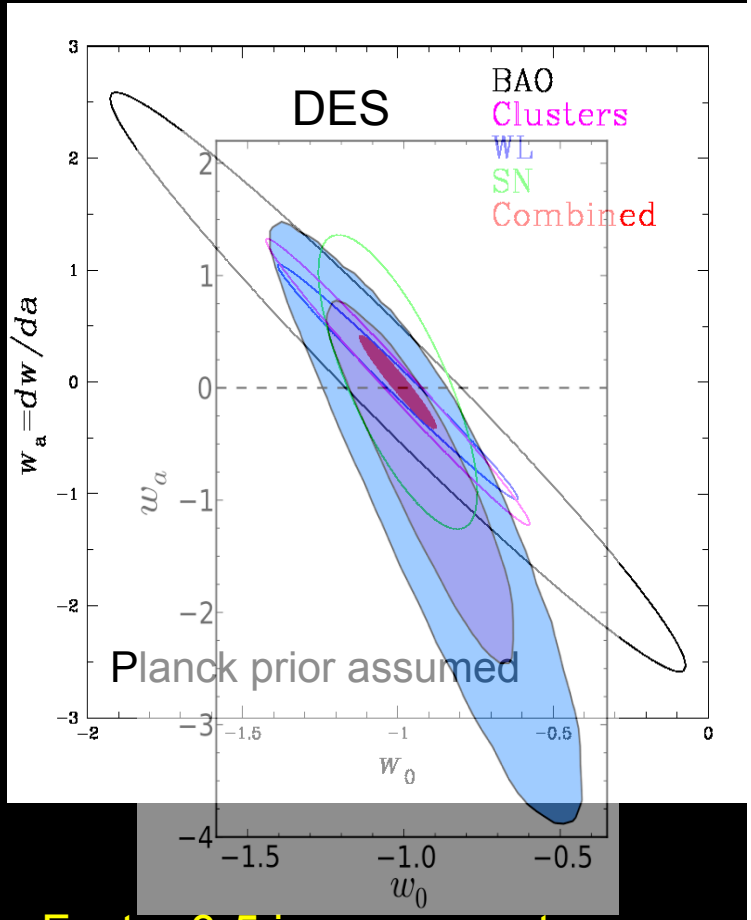
DES Science Summary

DARK ENERGY SURVEY

Four Probes of Dark Energy

- **Galaxy Clusters**
  - ~100,000 clusters to z>1
  - Synergy with SPT, VHS
  - Sensitive to growth of structure and geometry
- **Weak Lensing**
  - Shape measurements of 300 million galaxies
  - Sensitive to growth of structure and geometry
- **Baryon Acoustic Oscillations**
  - 300 million galaxies to z = 1 and beyond
  - Sensitive to geometry
- **Supernovae**
  - 30 sq deg time-domain survey
  - ~4000 well-sampled SNe Ia to z ~1
  - Sensitive to geometry

Forecast Constraints on DE Equation of State



Factor 3-5 improvement over Stage II DETF Figure of Merit

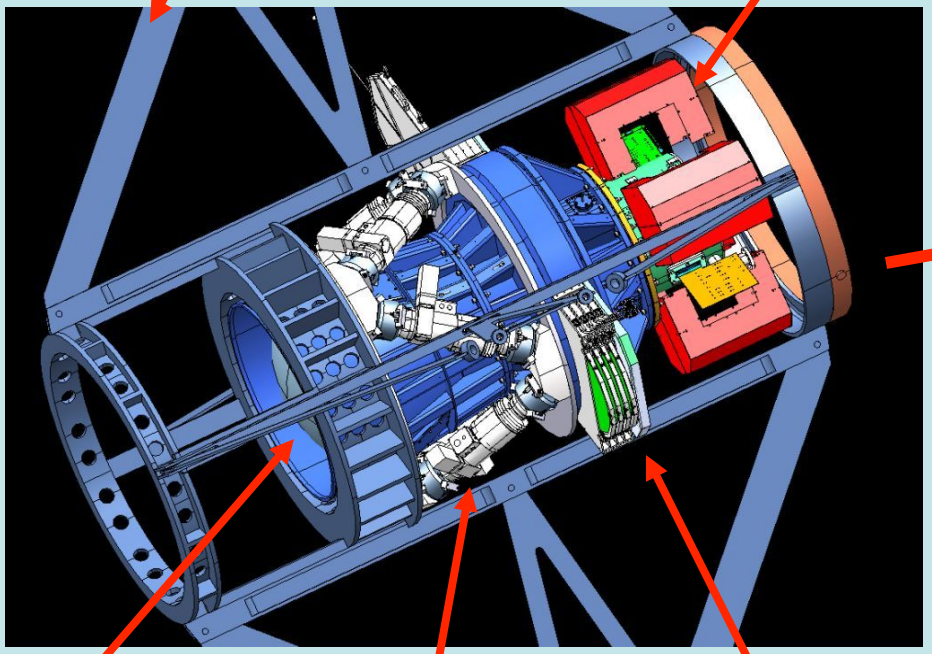
Project Structure & Timeline

- 3 Construction Projects:
  - DECam (hosted by FNAL; DOE supported)
  - Data Management System (NCSA; NSF support)
  - CTIO Facilities Improvement Project (NSF/NOAO)
    - NOAO Blanco Announcement of Opportunity 2003
    - DECam R&D 2004-8
    - Camera construction 2008-11
    - Final testing, integration now on-going
    - Shipping components to Chile this year
    - Installation on telescope ~Jan-May 2012
    - Commissioning and Science Verification: ~April-Aug. 2012
    - Survey operations begin: Sept 2012

Dark Energy Camera

Mechanical Interface of DECam Project to the Blanco

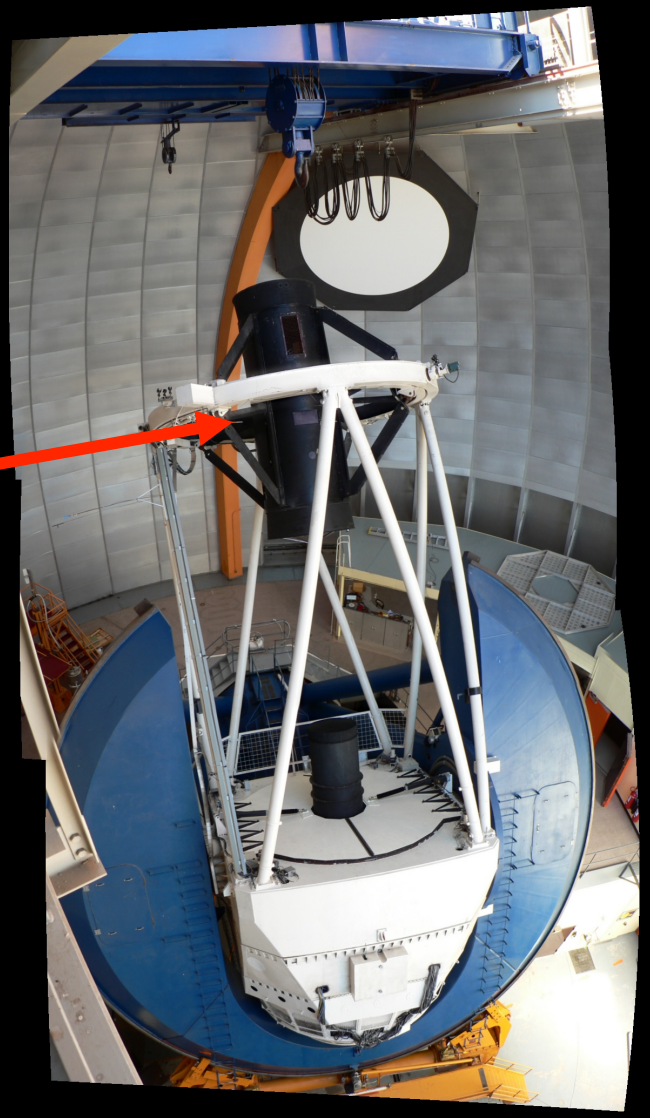
CCD Readout



Optical Corrector Lenses

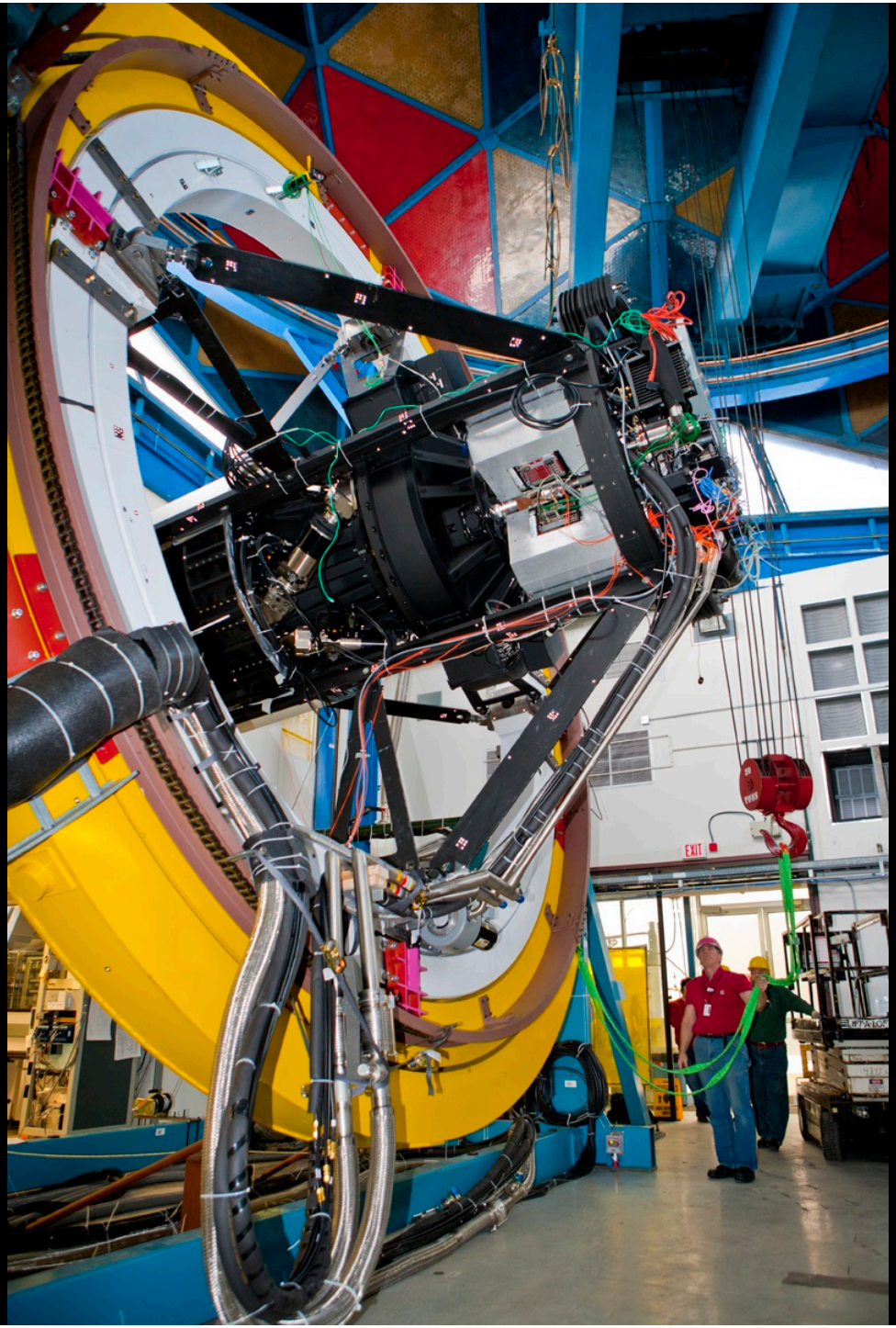
Hexapod: optical alignment

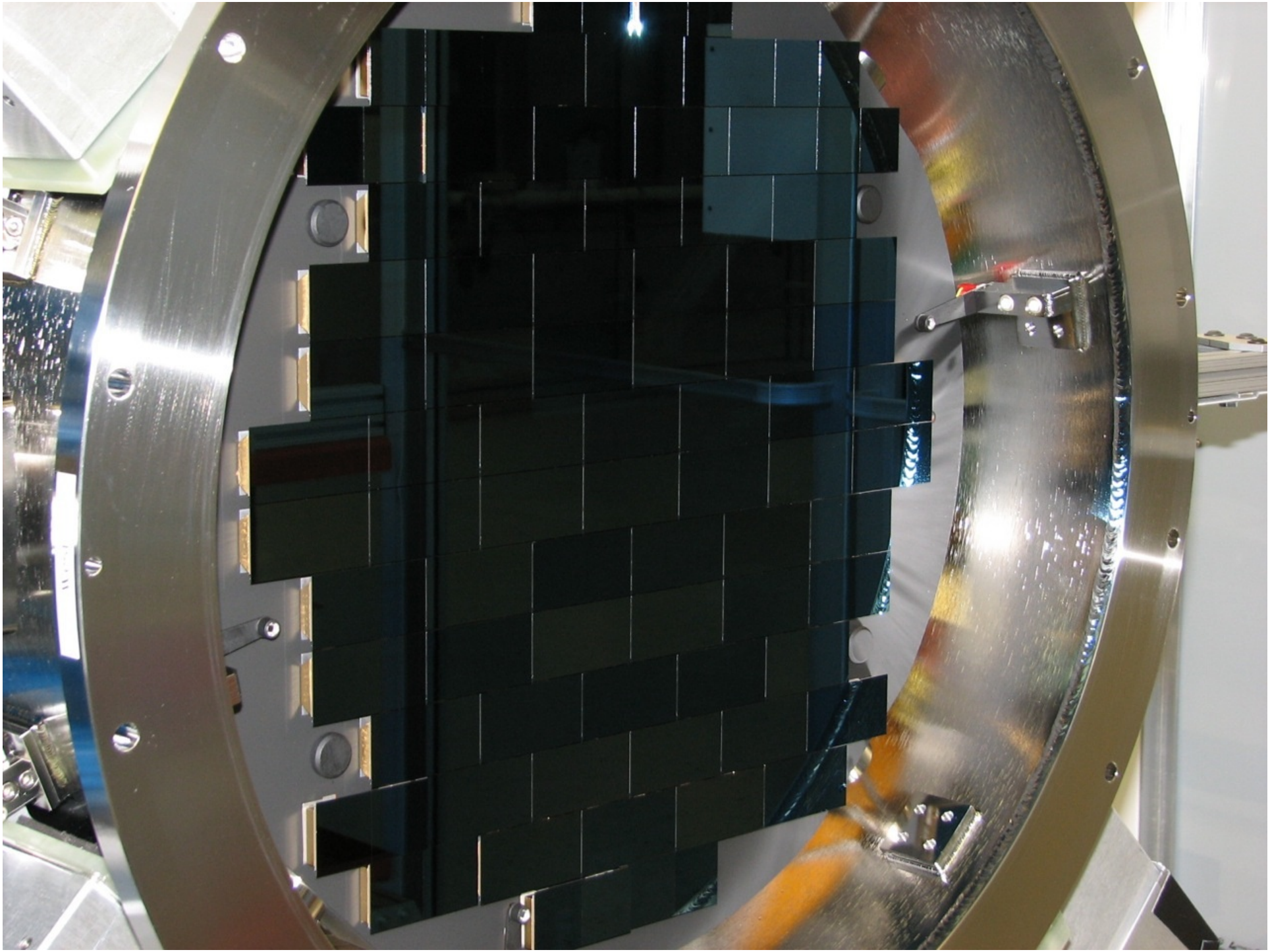
Filters & Shutter



DARK ENERGY SURVEY

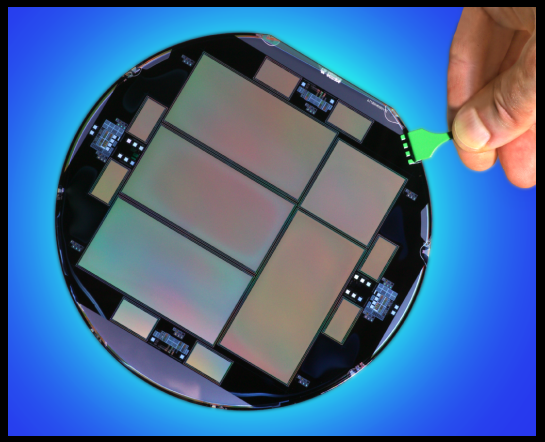
- DECam mounted on Telescope Simulator at Fermilab in early 2011
- DECam both DES survey instrument and CTIO facility instrument





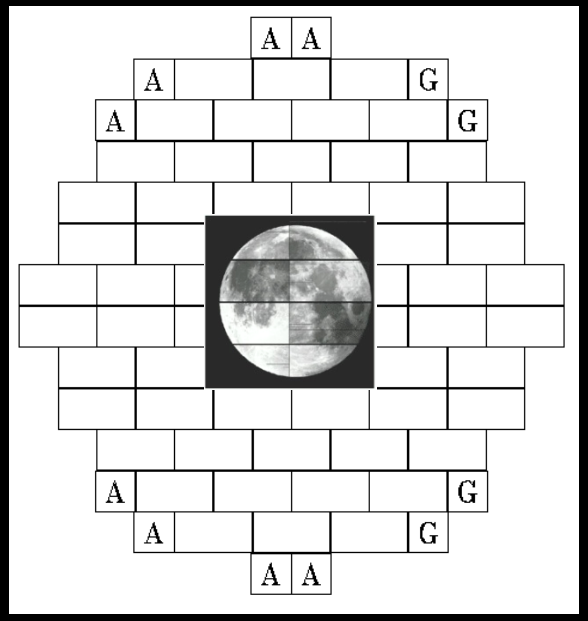
DARK ENERGY SURVEY

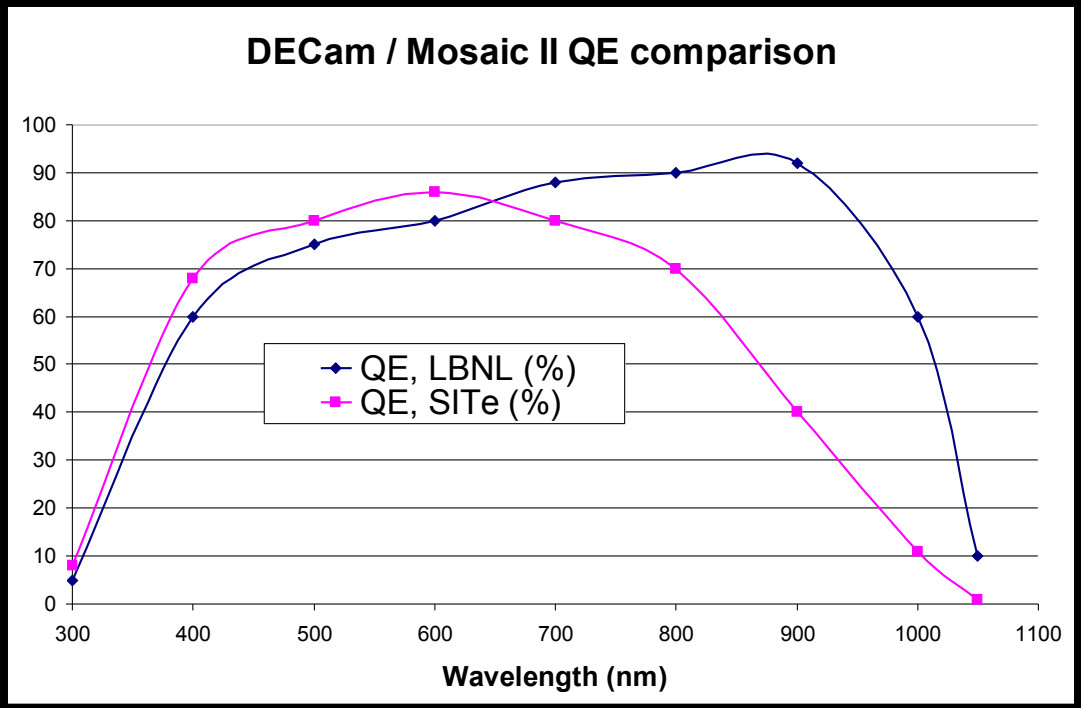
DECam CCDs



Developed by LBNL

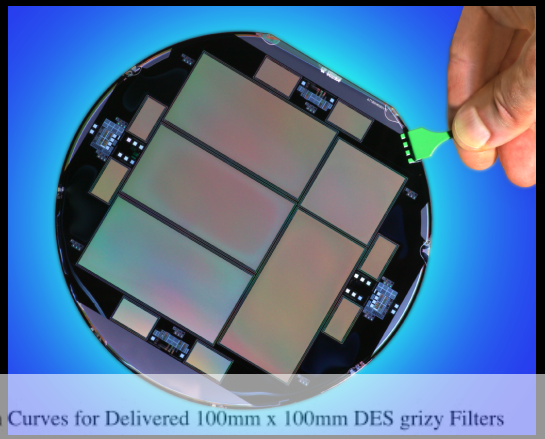
- 62 2kx4k fully depleted CCDs: 520 Megapixels, 250 micron thick, 15 micron (0.27”) pixel size
- 12 2kx2k guide and focus chips
- Excellent red sensitivity
- Roughly twice the number of science-grade CCDs packaged



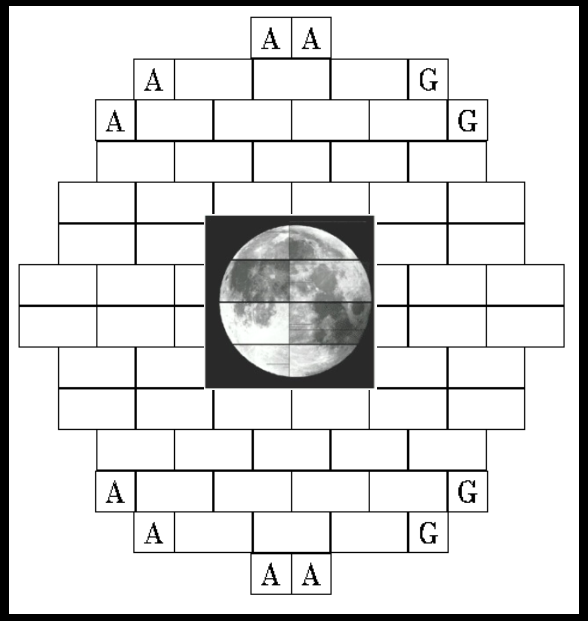


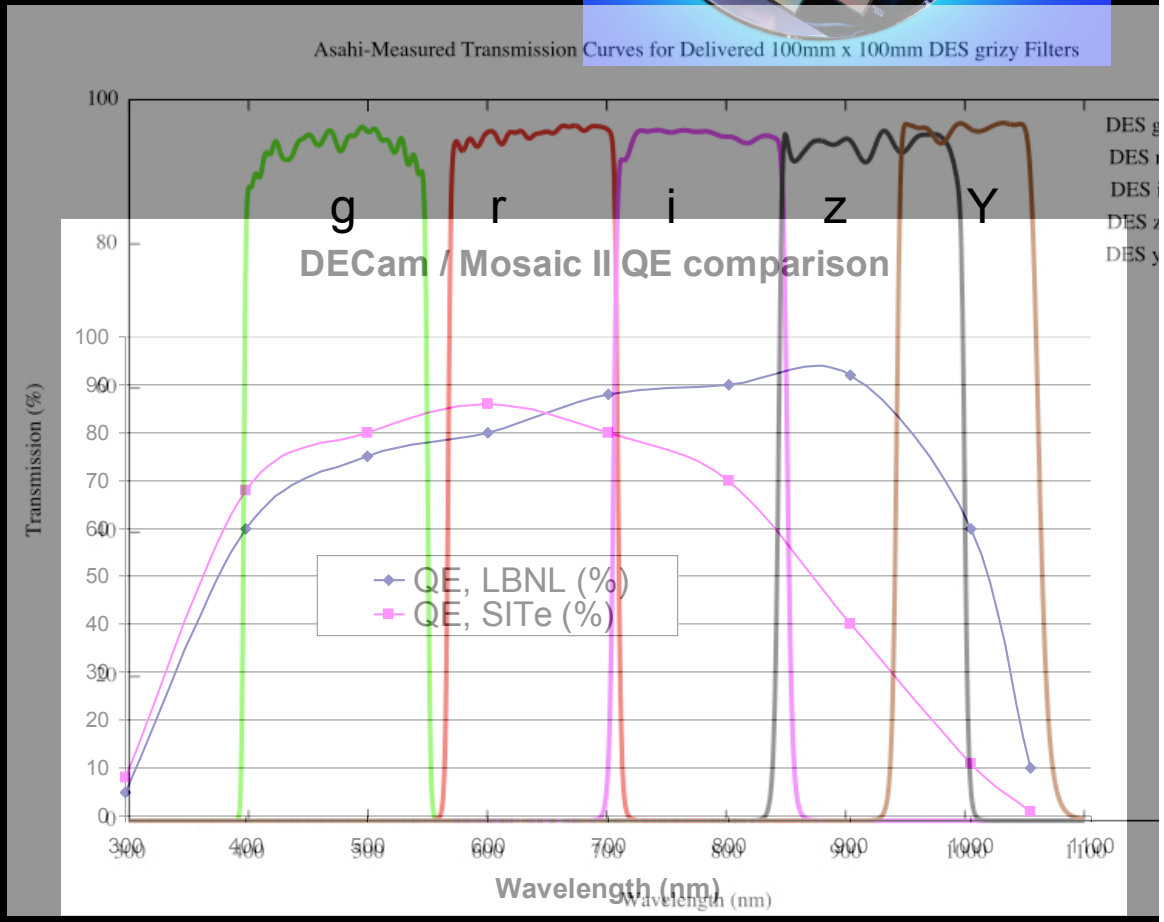
DARK ENERGY SURVEY

DECam CCDs



- 62 2kx4k fully depleted CCDs: 520 Megapixels, 250 micron thick, 15 micron (0.27”) pixel size
- 12 2kx2k guide and focus chips
- Excellent red sensitivity
- Roughly twice the number of science-grade CCDs packaged





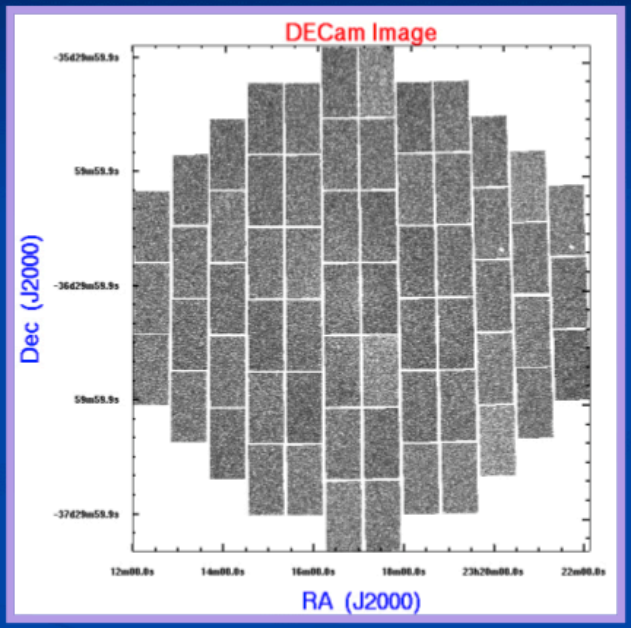
Asahi filters

DES Data Management

http://cosmology.illinois.edu/DES/

- The DESDM system:
  - Process DES data at NCSA
  - Archive DES data over the long term
    - ~4PB total, ~350TB database
  - Distribute data to Collaboration
    - Working groups analyze DM outputs
  - Distribute data to public
    - Raw/reduced data after 1 yr
    - Co-adds/catalogs at midpoint and end of survey

NOAO
NCSA

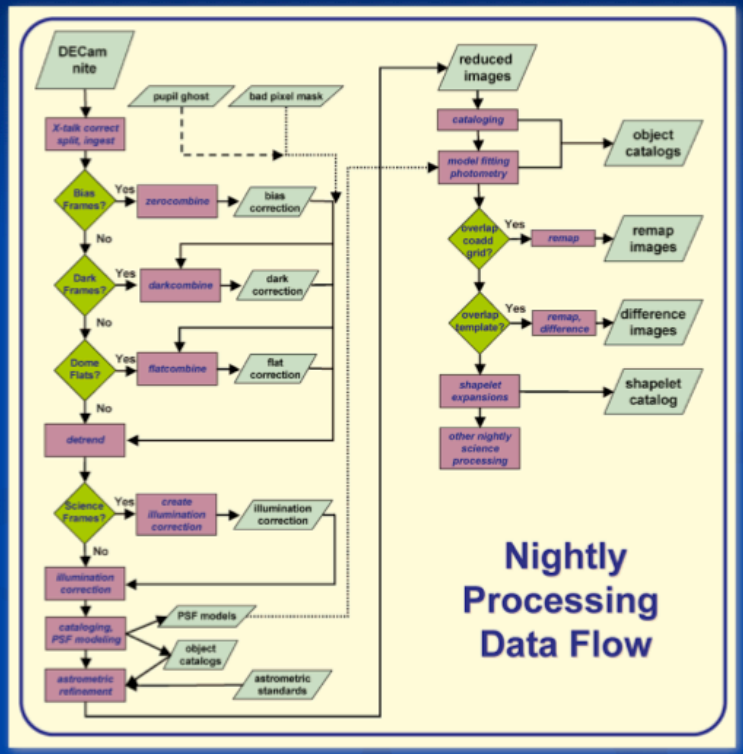


Exposure consists of 62 2kX4k CCD images - 3deg^2

Survey is ~150,000 exposures over 525 nights

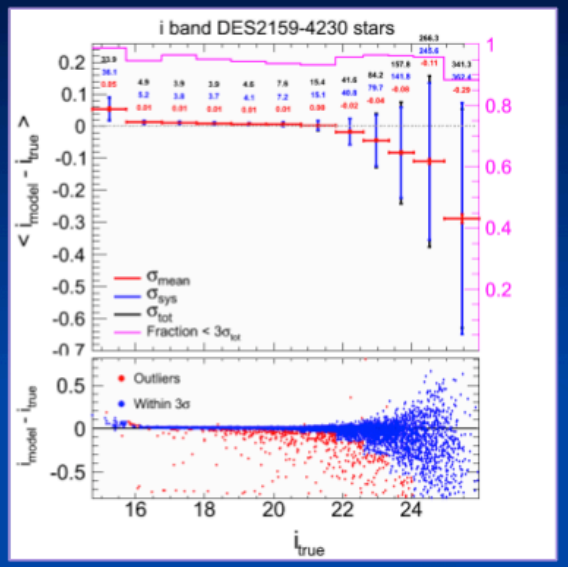
DESDM Processing

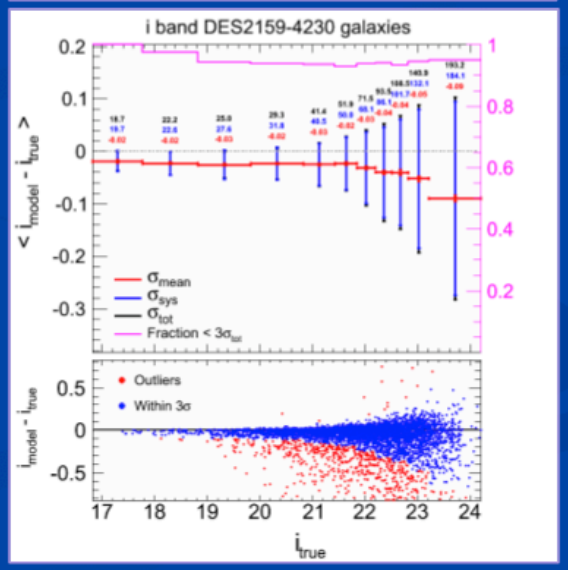
- DESDM processing includes:
  - Nightly processing
    - Detrending, astrometric refinement, remapping for coaddition and difference imaging, cataloging, ingestion to DES Archive and photometric calibration
  - Coaddition (w/PSF Homogenization)
    - Build and catalog deeper images of the sky
  - Weak Lensing
    - Extract shear measurements from the survey data (both single epoch and multi-epoch)
  - Difference Imaging
    - Support SNe science within dedicated fields
  - Photo-z, Survey Mask, etc



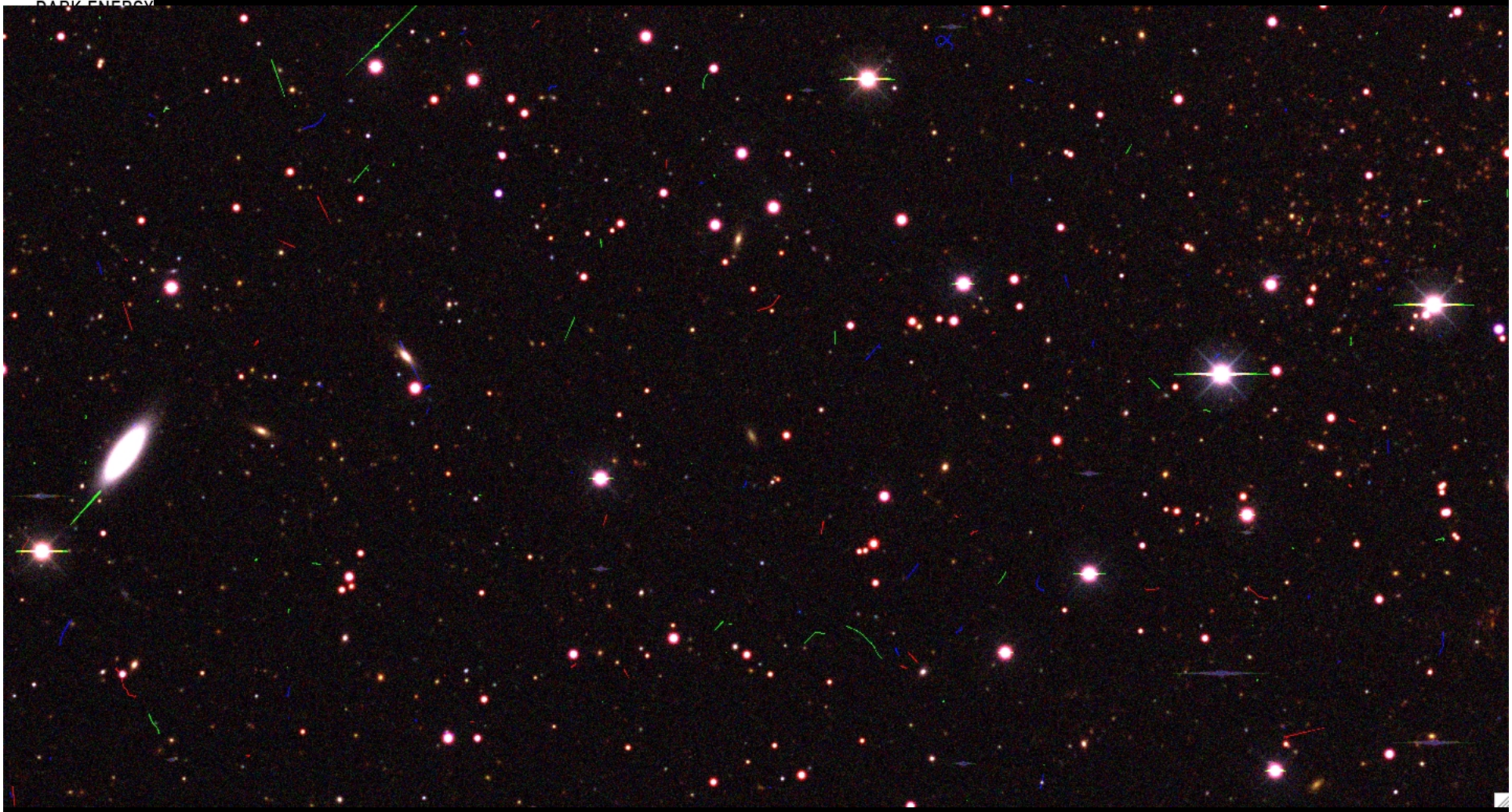
DES Data Quality

- Requirements:
  - <2% accurate photometric zeropoints
  - <100mas astrometry
  - Spatially uniform and accurate star-galaxy classification
- Strategy:
  - Develop DES specific detrending/calibration codes
  - Build upon existing AstroMatic toolkit from Bertin
- PSF Corrected Model Fitting Photometry
  - PSFEx used to model PSF variation across images
  - SExtractor extended to do PSF corrected model fitting
- PSF Homogenization for Coadds
  - Use PSFEx and tools that will be integrated into SWarp to build uniform median seeing coadds





DECam Image Simulations

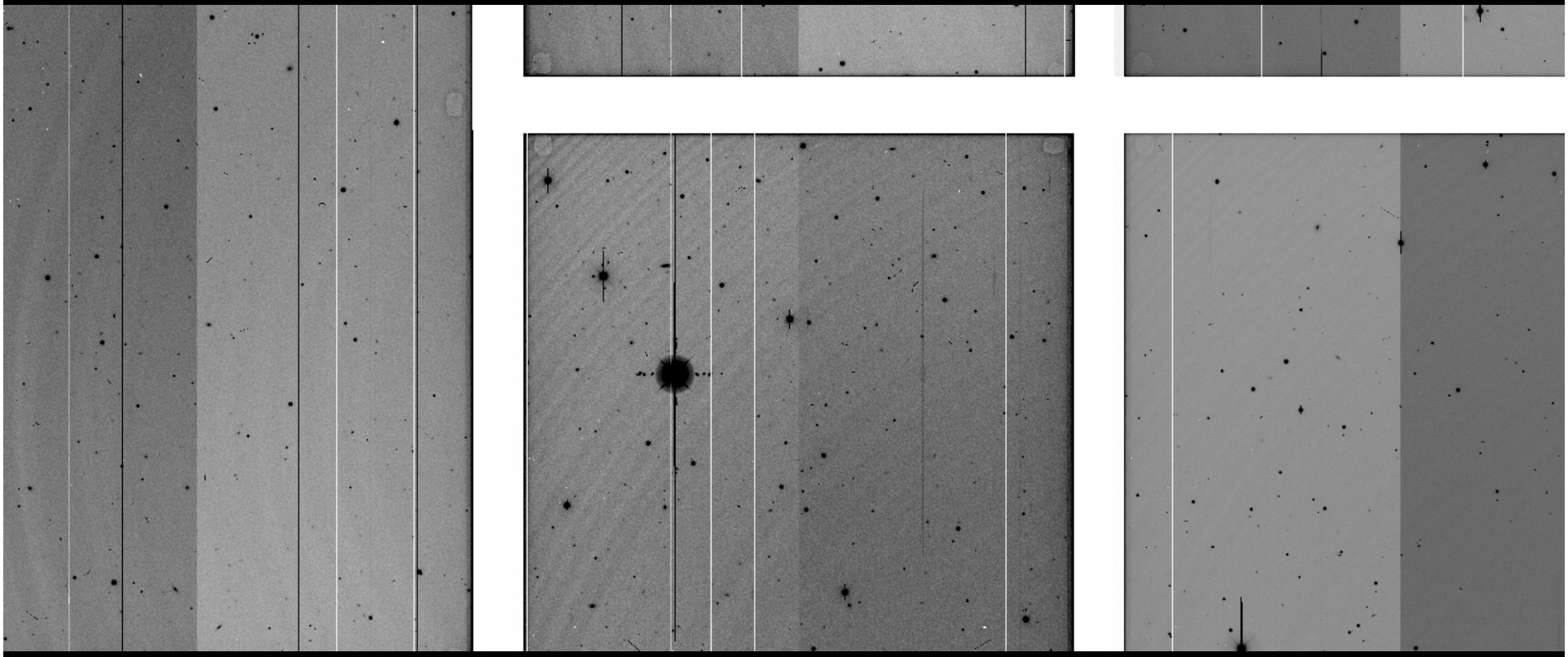


Populate N-body sims w/ galaxies drawn from SDSS+evolution+shapes

DECam Image Simulations

DARK ENERGY SURVEY

Series of Data Challenges to test Data Management System



- Note bright star artifacts, cosmic rays, cross talk, glowing edges, flatfield ("grind marks", tape bumps), bad columns, 2 amplifiers/CCD.
- Working groups analyze DM outputs → feedback to pipeline

Roles of Simulations

- Provide `truth' for testing data analysis & reduction software: e.g., cluster finding algorithms, shear and LSS statistics, photometric SN classification, Data Management testing.
- Central for data interpretation: e.g., mock catalogs for error covariances, emulation for parameter extraction, calibrating baryon physics impact on observables.
- Some of these are necessarily project-specific. Others benefit from a community development approach. Yet even project-specific sims could pay broader dividends by generalizing their frameworks.
- Exploring the boundary/relation between project- and community sim efforts would be beneficial to both.

Final Thoughts

- Data is messy compared to simulations (obvious).
- Data processing involves data modeling.
- Those analyzing public data often ignore this.
- It's not wholly their fault: packaging the metadata that quantifies messiness (selection functions, incompleteness, purity, e.g., deblending uncertainties, etc) can be challenging and often involves modeling decisions. Surveys should provide metadata needed to analyze or reproduce an analysis, but that often requires considerable analysis/modeling. Example: SN distance moduli.